



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/748,606	12/26/2000	Barbara A. Brown	APP1P010/44379/03338	4256

7590 07/02/2004

Steve Gupta  
Vice President Finance  
1100 Island Drive  
Redwood City, CA 94065

EXAMINER

FISHER, MICHAEL J

ART UNIT	PAPER NUMBER
----------	--------------

3629

DATE MAILED: 07/02/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

RECEIVED

JUL 21 2004

GROUP 3600

# Office Action Summary

Application No.

09/748,606

Applicant(s)

BROWN, BARBARA A.

Examiner

Michael J Fisher

Art Unit

3629

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

## Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

## Status

- 1) ☐ Responsive to communication(s) filed on \_\_\_\_.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

## Disposition of Claims

- 4) ☒ Claim(s) 1-18 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-18 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_ are subject to restriction and/or election requirement.

## Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

## Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some \* c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

## Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)  
Paper No(s)/Mail Date \_\_\_\_.
- 4) ☐ Interview Summary (PTO-413)  
Paper No(s)/Mail Date. \_\_\_\_.
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☐ Other: \_\_\_\_.

## **DETAILED ACTION**

### ***Claim Rejections - 35 USC § 101***

35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

Claims 1-4 and 6 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter. There is no technological innovation involved in the steps as they could be accomplished merely by talking between translators.

### ***Claim Rejections - 35 USC § 102***

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

Claims 1-18 are rejected under 35 U.S.C. 102(e) as being anticipated by US PAT 6,275,789 to Moser et al. (Moser).

As to claims 1,7 and 13, Moser discloses a computer apparatus for translation that receives a plurality of words for translation (data entry device 14), identifying the context associated with the words (col 35, lines 3-14), translating the words based on context (claim 2, section (f)).

As to claims 2,8 and 14, Moser discloses matching the words with a predetermined set of translated words (claim 1, section (d)).

As to claims 3,9 and 15, the predetermined words are selected based on context (claim 1, section (g)).

As to claims 4,6,10,12,16 and 17, the context is shown to be in a particular industry (col 34, lines 3-5 and further clarified in table 8 in col 20) and further, the words are shown to be technical words (table 8, such as words in "chemical/material" and "physics" and "botany").

As to claims 5,11 and 17, Moser discloses the device as using a network (fig 7).

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

Claims 1-4, 7-10 and 13-16 are rejected under 35 U.S.C. 102(b) as being anticipated by "Searchers Tap Into Dead Sea Scrolls Computer Said Able to Recast Documents" by Wilford (Wilford).

As to claim 1, the Dead Sea Scrolls were translated by receiving a plurality of words to be translated (those languages on the scrolls that were unknown), identifying a context associated with the words (they were all on the scrolls), translating the words based on the identified context (they were translated because they were on the scrolls).

Art Unit: 3629

As to claim 2, they were able to be translated because they were matched with a predetermined set of words (those on the scrolls that were in a known language).

As to claim 3, they were selected based on the identified context.

As to claim 4, the context included a particular industry (archaeology).

As to claims 7-10 and 13-16, as Wilford discloses translating the scrolls using a computer, the use of computer codes and computer logic would be used in the translation process.

### ***Claim Rejections - 35 USC § 103***

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

The factual inquiries set forth in *Graham v. John Deere Co.*, 383 U.S. 1, 148 USPQ 459 (1966), that are applied for establishing a background for determining obviousness under 35 U.S.C. 103(a) are summarized as follows:

1. Determining the scope and contents of the prior art.
2. Ascertaining the differences between the prior art and the claims at issue.
3. Resolving the level of ordinary skill in the pertinent art.
4. Considering objective evidence present in the application indicating obviousness or nonobviousness.

Claims 5, 11 and 17 are rejected under 35 U.S.C. 103(a) as being unpatentable over Moser. Moser discloses a method of translation as discussed above. Moser does not, however, teach receiving the words via a network. As Moser uses a computer and

Art Unit: 3629

further, as it is very well known in the art to connect a computer to the Internet, it would have been obvious to one of ordinary skill in the art to modify the system as disclosed by Moser by receiving the words over a network to try to ensure that the proper words are inputted as receiving them over a network would only require them to be entered once instead of each time someone tries to do a translation.

### ***Conclusion***

The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. US PAT 6,546,365 to Gajda et al., Gajda et al. disclose a method and system for translation using a table with translated words (title), Japanese patent JP10031672A to NEC, NEC discloses a translation program that uses a particular context (names) and further includes a predetermined set of translated words (the names).

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Michael J Fisher whose telephone number is 703-306-5993. The examiner can normally be reached on Mon.-Fri. 7:30am-5:00pm alt Fri. off.

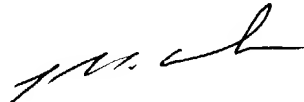
The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Art Unit: 3629

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

MF 

6/27/04



JOHN G. WEISS  
SUPERVISORY PATENT EXAMINER  
TECHNOLOGY CENTER 3600

<b>Notice of References Cited</b>	Application/Control No. 09/748,606	Applicant(s)/Patent Under Reexamination BROWN, BARBARA A.	
	Examiner Michael J Fisher	Art Unit 3629	Page 1 of 1

#### U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
	A	US-6,546,365	04-2003	Gajda et al.	704/8
	B	US-6,275,789	08-2001	Moser et al.	704/7
	C	US-			
	D	US-			
	E	US-			
	F	US-			
	G	US-			
	H	US-			
	I	US-			
	J	US-			
	K	US-			
	L	US-			
	M	US-			


#### FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N	JP 10031672 A	02-1998	Japan	NEC Home Electronics	G06F 17/22
	O					
	P					
	Q					
	R					
	S					
	T					

#### NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	"Researchers Tap Into Dead Sea Scrolls Computer Said Able To Recast Documents" by Wilford, 9/1991, Phoenix Gazette, P A3, entire document.
	V	
	W	
	X	

\*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)  
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

 6/27/04



DERWENT-ACC-NO: 1998-165152

DERWENT-WEEK: 199815

COPYRIGHT 1999 DERWENT INFORMATION LTD

TITLE: Name translator of Japanese language word processor, PC  
- includes comparison unit which outputs person name from  
which Chinese characters other than person name notation  
are deleted

PATENT-ASSIGNEE: NEC HOME ELECTRONICS LTD[NIDF]

PRIORITY-DATA: 1996JP-0205309 (July 16, 1996)

PATENT-FAMILY:

PUB-NO	PUB-DATE	LANGUAGE	PAGES	MAIN-IPC
JP 10031672 A	February 3, 1998	N/A	005	G06F 017/22

APPLICATION-DATA:

PUB-NO	APPL-DESCRIPTOR	APPL-NO	APPL-DATE
JP 10031672A	N/A	1996JP-0205309	July 16, 1996

INT-CL (IPC): G06F017/22

ABSTRACTED-PUB-NO: JP 10031672A

BASIC-ABSTRACT:

The converter includes an input unit (1) which inputs a person name. An analysis unit (21) divides the Kana character row of the input person name and analyses the obtained division character string. A single Chinese character dictionary (25) stores single Chinese character corresponding to the analysed division character string. An exceptional dictionary (26) stores Chinese character other than person name notation. A search unit (22) searches for the single Chinese character corresponding to the division character string, from the single Chinese character dictionary.

A synthesising unit (23) synthesises the person name from the searched single Chinese character and the row of the division character string. A comparison

unit (24) compares the synthesised person name with the exceptional dictionary and deletes the Chinese character other than the name notation. The comparison unit has an output part which outputs the person name from which Chinese characters other than person name notation are deleted.

ADVANTAGE - Enables to perform various person name conversion process appropriately.

CHOSEN-DRAWING: Dwg.1/5

DERWENT-CLASS: T01

EPI-CODES: T01-J11A;

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-31672

(43) 公開日 平成10年(1998) 2月3日

(51) Int.Cl.<sup>6</sup>

G 0 6 F 17/22

識別記号

庁内整理番号

F I

G 0 6 F 15/20

技術表示箇所

5 2 0 L

審査請求 未請求 請求項の数 4 F D (全 5 頁)

(21) 出願番号 特願平8-205309

(22) 出願日 平成8年(1996) 7月16日

(71) 出願人 000001937

日本電気ホームエレクトロニクス株式会社  
大阪府大阪市中央区城見一丁目4番24号

(72) 発明者 本間 一哉

大阪府大阪市中央区城見一丁目4番24号  
日本電気ホームエレクトロニクス株式会社  
内

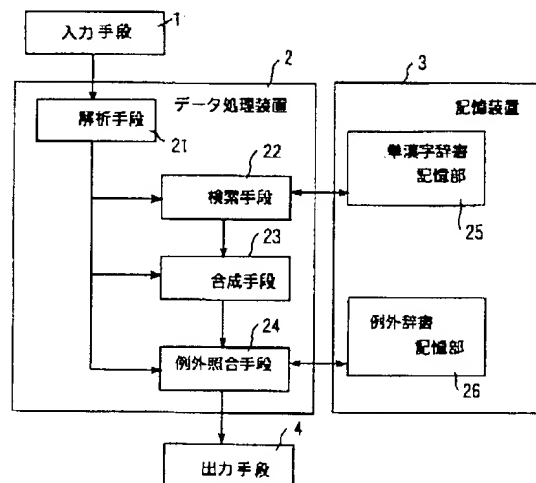
(74) 代理人 弁理士 飯塚 信市

(54) 【発明の名称】 人名変換装置及び方法

(57) 【要約】

【課題】 人名用に膨大な漢字候補テーブルを構築することなく、この種の人名変換処理を適切に行い得るようにした人名変換装置及び方法を提供すること。

【解決する手段】 入力手段から入力された人名の仮名文字列を分割し、この分割された分割文字列で読みを解析する解析手段と、この解析手段で解析された分割文字列の読みに対応する単漢字を記憶する単漢字辞書と、上記人名表記以外の漢字を記憶する例外辞書と、上記単漢字辞書に記憶された単漢字から上記分割文字列に対応する単漢字を検索する検索手段と、この検索手段で検索された単漢字と上記分割文字列の並びとから人名を合成する合成手段と、この合成手段で合成された人名を上記例外辞書と照合して上記人名表記以外の漢字を削除する照合手段と、この照合手段が上記人名表記以外の漢字を削除して得た人名を出力する出力手段とで構成される。



## 【特許請求の範囲】

【請求項1】 入力手段から入力された人名の仮名文字列を分割し、この分割された分割文字列で読みを解析する解析手段と、この解析手段で解析された分割文字列の読みに対応する単漢字を記憶する単漢字辞書と、上記人名表記以外の漢字を記憶する例外辞書と、上記単漢字辞書に記憶された単漢字から上記分割文字列に対応する単漢字を検索する検索手段と、この検索手段で検索された単漢字と上記分割文字列の並びとから人名を合成する合成手段と、この合成手段で合成された人名を上記例外辞書と照合して上記人名表記以外の漢字を削除する照合手段と、この照合手段が上記人名表記以外の漢字を削除して得た人名を出力する出力手段とで構成されたことを特徴とする人名変換装置。

【請求項2】 上記分割文字列に対応する読みを見出しとする見出し語と、この見出し語が上記分割文字列のどの位置に置かれるかを設定した位置情報と、この位置情報で設定される上記単漢字とを記憶する上記単漢字辞書で構成され、上記位置情報によって上記分割文字列に対応する上記単漢字を検索することを特徴とする請求項1に記載の人名変換装置。

【請求項3】 入力手段から入力された人名の仮名文字列を分割し、この分割された分割文字列で読みを解析手段が解析し、この解析された分割文字列の読みに対応する単漢字を単漢字辞書から検索手段に検索させ、この検索手段で検索された単漢字と上記分割文字列の並びとから人名を合成手段が合成し、この合成された人名を例外辞書と照合して上記人名表記以外の漢字を照合手段で削除させ、この照合手段が上記人名表記以外の漢字を削除して得た人名を出力手段で出力することを特徴とする人名変換方法。

【請求項4】 上記解析された分割文字列に対応する読みを見出し語から検索し、この見出し語の位置情報によって上記分割文字列に対応する上記単漢字を検索することを特徴とする請求項3に記載の人名変換方法。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】 本発明は、日本語ワードプロセッサ、およびそれに類するソフトウェアプログラムに於いて機能し、特に人名の読みを入力した場合における漢字候補の表示を自動的に行なう人名変換装置及び方法に関する。

## 【0002】

【従来の技術】 従来、日本語ワードプロセッサあるいはパーソナルコンピュータ上で機能する同様のソフトウェアプログラムを用いて、入力された読みを漢字に変換する場合には、所謂辞書機能によって読みに対する漢字候補を順次、あるいは一覧表示し、これらの候補の中から使用者が適切な漢字候補を選択することが行なわれている。かかる辞書機能は、メモリ内に格納された漢字候補

テーブルを参照して読みに対応する漢字候補が選択される。

【0003】 この漢字候補テーブルは、通常の辞書に記載されているような一般用語の他、地名、人名等の固有名詞をも格納しておくことが一般的である。また、長文の読みが入力された場合には、この長文を文節毎に区切る文節判断機能と共働させて文節ごとに対応する漢字候補を表示することができるようになっている。

## 【0004】

10 【発明が解決しようとする課題】 しかしながら文中に現われる固有名詞、中でも人名は、一般名詞以上に様々な読み、様々な候補があるため、辞書機能を十分機能させるためには、人名用に膨大な漢字候補テーブルを構築する必要があった。

【0005】 例えば同じ「かずや」と言う読みの人名であっても「一也」「一弥」「和也」「一哉」等の様々な漢字候補が存在する一方で、例えば「数屋」なる名は不適切な候補であると言える。従って、これらの全てを漢字候補として表示するためには漢字候補テーブルにその分だけ予め辞書登録しておかねばならず、漢字候補テーブルの格納されるメモリの記憶容量が増大するという課題を有していた。

【0006】 この発明は、従来の人名変換処理における以上の問題点を鑑みてなされたものであり、その目的とするところは、人名用に膨大な漢字候補テーブルを構築することなく、この種の人名変換処理を適切に行い得るようにした人名変換装置及び方法を提供することにある。

## 【0007】

30 【課題を解決するための手段】 この出願の請求項1に記載の発明は、入力手段から入力された人名の仮名文字列を分割し、この分割された分割文字列で読みを解析する解析手段と、この解析手段で解析された分割文字列の読みに対応する単漢字を記憶する単漢字辞書と、上記人名表記以外の漢字を記憶する例外辞書と、上記単漢字辞書に記憶された単漢字から上記分割文字列に対応する単漢字を検索する検索手段と、この検索手段で検索された単漢字と上記分割文字列の並びとから人名を合成する合成手段と、この合成手段で合成された人名を上記例外辞書と照合して上記人名表記以外の漢字を削除する照合手段と、この照合手段が上記人名表記以外の漢字を削除して得た人名を出力する出力手段とで構成されたことを特徴とする人名変換装置にある。

【0008】 そして、この請求項1に記載の発明によれば、入力手段から入力された人名の仮名文字列の分割された分割文字列で読みが解析され、この解析手段で解析された分割文字列の読みに対応する単漢字を記憶する単漢字辞書に記憶された単漢字から上記分割文字列に対応する単漢字が検索され、この検索された単漢字と上記分割文字列の並びとから人名が合成され、この合成された

人名は人名表記以外の漢字を記憶する例外辞書と照合して人名表記以外の漢字が削除され、その結果最終的に得られた人名が出力されることとなる。

【0009】また、この出願の請求項2に記載の発明は、上記分割文字列に対応する読みを見出しとする見出し語と、この見出し語が上記分割文字列のどの位置に置かれるかを設定した位置情報と、この位置情報で設定される上記単漢字とを記憶する上記単漢字辞書で構成され、上記位置情報によって上記分割文字列に対応する上記単漢字を検索することを特徴とする請求項1に記載の人名変換装置にある。

【0010】そして、この請求項2に記載の発明によれば、請求項1の構成において、上記分割文字列に対応する上記単漢字は、上記分割文字列に対応する読みを見出しとする見出し語が上記分割文字列のどの位置に置かれるかを設定した位置情報に基づいて検索される。

【0011】また、この出願の請求項3に記載の発明は、入力手段から入力された人名の仮名文字列を分割し、この分割された分割文字列で読みを解析手段が解析し、この解析された分割文字列の読みに対応する単漢字を単漢字辞書から検索手段に検索させ、この検索手段で検索された単漢字と上記分割文字列の並びとから人名を合成手段が合成し、この合成された人名を例外辞書と照合して上記人名表記以外の漢字を照合手段で削除させ、この照合手段が上記人名表記以外の漢字を削除して得た人名を出力手段で出力することを特徴とする人名変換方法にある。

【0012】また、この出願の請求項4に記載の発明は、上記解析された分割文字列に対応する読みを見出し語から検索し、この見出し語の位置情報によって上記分割文字列に対応する上記単漢字を検索することを特徴とする請求項3に記載の人名変換方法にある。

【0013】そして、以上の各請求項の発明によれば、人名用に膨大な漢字候補テーブルを構築することなく、この種の人名変換処理を適切に行なうことができる。

【0014】

【発明の実施の形態】以下、本発明の好適な実施の形態を図面を参照して詳細に説明する。

【0015】図1は本発明にかかる人名変換装置の一実施形態を示すブロック図である。同図に示されるように、この人名変換装置は、主としてキーボード等の入力手段1と、プログラム制御により動作するデータ処理装置2と、情報を記憶する記憶装置3と、ディスプレイ装置等の出力手段4とから構成されている。

【0016】入力手段1は、仮名、ローマ字等により読みを入力するものである。データ処理装置2は入力手段1から入力された読みに対応した仮名文字列を便宜的に部分文字列に分割する解析手段21と、解析手段21により分割された仮名文字列から後述する単漢字辞書記憶部25を参照して対応する単漢字を検索する検索手段2

2と、検索された単漢字を読みに対応して合成する合成手段23と、合成された漢字列から後述する例外辞書記憶部26を参照して不適切な漢字列を除去する照合手段24を含んでいる。

【0017】記憶装置3は、人名に使用される頻度の高い単漢字を記憶するとともに、その単漢字が使用される可能性のある人名の読みの中での位置情報を示す位置フラグをも記憶する単漢字辞書記憶部25と、ある読みに対して使用されることがないと思われる単漢字の組合せを記憶してある例外辞書記憶部26とからなる。

【0018】出力手段4は入力手段1から入力された読み入力、データ処理装置2によって処理されて検索された漢字候補群等をディスプレイ表示する。

【0019】次に、単漢字辞書記憶部25及び例外辞書記憶部26の一実施形態の詳細について図2及び図3を参照して説明する。

【0020】図2は単漢字辞書記憶部25に格納されている辞書情報の一実施例を示すテーブルであり、このテーブルには見出し語311と、漢字列内での位置情報を示す位置フラグ312と、漢字候補313とが予め記憶してある。

【0021】ここで見出し語311は、漢字候補のそれぞれに対応する読みを示すものであり、位置フラグ312はその読みが、例えば人名において先頭に位置する可能性があるのか、中間に位置する可能性があるのか、あるいは末尾に位置する可能性があるのかを示すものである。例えば「かず」と言う読みは、「かずや」「かずこ」「かずお」等のように、名前の先頭に位置する可能性と、「しげかず」「ともかず」等のように名前の末尾に位置する可能性がある。一方、名前の中間に位置することはほとんどない。位置フラグは、このように慣習的に定めた位置情報を3桁の情報として示すものであり、例えば「かず」の場合には、名前の先頭と末尾に位置する可能性があるので「101」という位置フラグとなる。

【0022】漢字候補313は、見出し語311とは一対多の関係にあり、同じ「かず」の読みに対して「一」「和」等の単漢字がそれぞれ対応する。

【0023】一方例外辞書記憶部26は、図3に示すように見出し語321と、その見出し語に対応する漢字列としては不適切な人名322を一对一、あるいは一対多の関係にて対応させたテーブルである。

【0024】次に本発明にかかる人名変換装置の作用を図4及び図5を適宜参照しながら説明する。

【0025】まず、解析手段21は、入力手段1から入力された仮名文字列を分割するための区切り情報を付加し、区切り情報の付加された仮名文字列は部分文字列の並びとして漢字検索手段22に供給される。

【0026】例えば「かずや」という仮名文字列が入力されると、読み解析手段としては「かずや」「か／ず

や」「かず／や」「か／ず／や」のように仮名文字列を分割する。なお「／」は区切り情報を示すものであり、この区切り情報で区切られた仮名文字あるいは仮名文字列が部分文字列となる。

【0027】解析手段により得られた分割文字列は、それぞれ漢字検索手段22によって単漢字辞書記憶部25を参照して、該当する単漢字候補が検索される。図4は検索手段22の行なう処理を示すフローチャートである。

【0028】まず解析手段21で分割された分割文字列のその読み全体の中での位置情報を検索フラグとして設定する(ステップ401)。この時、例えば「かず」と言う分割文字列を検索しようとする場合、「かず」は、「かず／や」なる仮名文字列の先頭に位置するものであるので、検索用の位置フラグとして「100」を設定する。そして「かず」という読みを検索時のキー項目として設定し(ステップ402)、このキー項目に対応した見出し語及び単漢字を検索する(ステップ403)。もしも対応する見出し語及び単漢字が単漢字辞書記憶部25に存在するかどうか判断され(ステップ404)、該当する単漢字が存在する場合には、検索フラグと単漢字辞書記憶部25に記憶された位置フラグとが一致するかどうかさらに判断される(ステップ405)。これらの判断を全て満足する場合には、検索された単漢字を候補の一つとして記憶する(ステップ406)。

【0029】ステップ403からステップ406までの処理は、部分文字、及びその検索フラグに対応する単漢字がすべて単漢字辞書記憶部25から検索されるまで繰り返される。この結果、例えば「かず」という部分文字に対しては「一」「和」が検索され「や」という部分文字に対しては「哉」「也」「弥」「矢」「八」等の単漢字候補が検索される。

【0030】なお、もしも「やいち」なる読みが入力され、これを部分文字列に分割すると「やいち」「や／いち」「やい／ち」「や／い／ち」のようになるが、この中で「やい／ち」に該当する単漢字列を検索しようすると、「やい」なる単漢字候補は存在しないので、誤った単漢字候補は検索されない。また「や／いち」に該当する単漢字列を検索しようとする場合、「や」は先頭に位置するため検索フラグは「100」となるが、同じ見出し語「や」を有する単漢字候補のうち「也」「哉」は検索フラグと一致しないので、これらの単漢字候補は検索されない。

【0031】これらの処理を終了したならば、検索された単漢字は次の合成手段23に送られる。

【0032】合成手段23は検索された単漢字を、位置フラグと対応させながら合成する。例えば「かず／や」の場合、先に検索された単漢字候補の組合せから「一哉」「一也」「一弥」「一矢」「一八」「和哉」「和也」「和弥」「和矢」「和八」が合成手段により組み合

わされる。

【0033】次に合成手段23により組み合わせられた候補の中から不適切な名を除外するために例外照合手段24および例外辞書記憶部26が用いられる。図5は、照合手段24の行なう処理を示すフローチャートである。まず照合手段24は与えられた読みに対する不適切な名があるかどうかを例外辞書記憶部26から検索し(ステップ501)、合成手段23により合成された名と、検索された不適切な名とを照合し(ステップ502)、その不適切な名に一致しているかどうかを判断する(ステップ503)。もしも一致する場合、その合成文字列候補を除外する(ステップ504)。本実施形態の場合、例外辞書記憶部26には「かずや」なる読みに対して「和八」なる漢字列は不適切なものであることが記憶されているので、合成された漢字列から「和八」が除去される。

【0034】以上のようにして、照合手段24を経て得られる候補は、入力された読みに対する適切な人名の候補群となり、これらを順次あるいは一括表示して使用者の選択を待つことになる。

【0035】

【発明の効果】以上の説明で明らかなように、この発明によれば、入力手段から入力された人名の仮名文字列を分割し、この分割された分割文字列で読みが解析され、この解析された分割文字列の読みに対応する単漢字が単漢字辞書から検索され、この検索された単漢字と分割文字列の並びとから人名が合成され、この合成された人名を例外辞書と照合することにより人名表記以外の漢字が削除され、この人名表記以外の漢字を削除して得た人名が出力されることとなり、その結果、人名用に膨大な漢字候補テーブルを構築することなく、この種の人名変換処理を適切に行なうことができる。

【図面の簡単な説明】

【図1】本発明にかかる人名の変換装置の一実施例を示すブロック図である。

【図2】図1の人名変換装置における単漢字辞書記憶部25の内部を説明するための図である。

【図3】図1の人名変換装置における例外辞書記憶部32の内部を説明するための図である。

【図4】図1の人名変換装置における漢字検索手段22の行なう処理を示すフローチャートである。

【図5】図1の人名変換装置における例外照合処理手段24の行なう処理を示すフローチャートである。

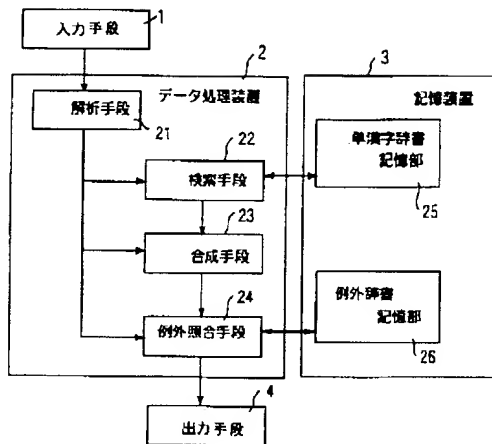
【符号の説明】

- 1 入力手段
- 2 データ処理装置
- 3 記憶装置
- 4 出力手段
- 21 解析手段
- 22 検索手段

23 合成手段  
24 例外照合手段

25 単漢字辞書記憶部  
26 例外辞書記憶部

【図1】



【図2】

見出し語	位置情報	単漢字
かず	1 0 1	和
かず	1 0 1	裁
や	0 1 1	也
や	0 1 1	矢
や	1 1 1	八
や	1 1 1	ハ
や	1 0 1	

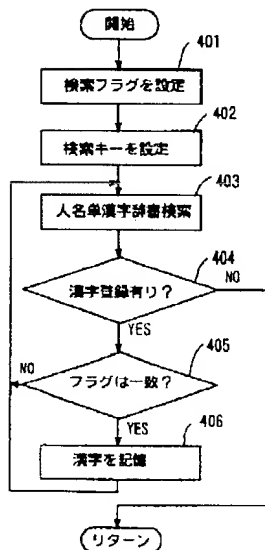
311 312 313

【図3】

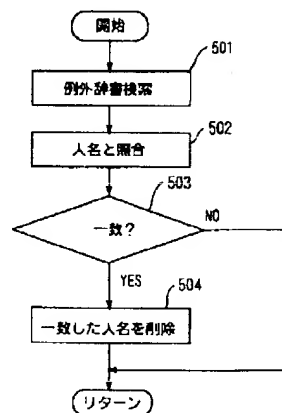
見出し語	不適切な人名
かずや	和八

321 322

【図4】



【図5】



# RESEARCHERS TAP INTO DEAD SEA SCROLLS COMPUTER SAID ABLE TO RECAST DOCUMENTS

Phoenix Gazette (PG) - THURSDAY September 5, 1991

By: John Noble Wilford, The New York Times

Edition: STREET FINAL Section: Front Page: A3

Word Count: 570

## TEXT:

NEW YORK - Researchers using a desktop computer say they have broken the grip a coterie of scholars has held for 40 years on the **Dead Sea Scrolls**, one of the great finds of biblical archaeology.

The scholars have published transcriptions of less than half of the scrolls and have allowed only limited access to the rest. The scrolls, containing accounts and insights into the history of Judaism and the origins of Christianity, have long tantalized researchers who have sharply criticized the slow pace in publishing transcriptions and **translations** of them.

But in what amounts to an end run around the scholarly blockade, two researchers at Hebrew Union College in Cincinnati devised a computer program that used a listing of all the words in one collection of scrolls to reconstruct part of the original text.

The listing, called a concordance, is organized like a dictionary, listing every use of every word that appears in a text, identifying the places where the word is found and also giving the context in which the word is used. Similar concordances have been prepared for most of the texts found in 11 caves near the Dead Sea in 1947, and the computer method could theoretically be used to produce texts from other unpublished scroll material as well.

In announcing the publication at a news conference in New York, Hershel Shanks, president of the Biblical Archaeology Society, said the first volume contained no historical or theological "bombshells." Instead, he said, its significance lay in the publication itself.

"We've broken the monopoly," he said, referring to the small group of editors who have controlled the study and publication of scroll material over the last 40 years.

These editors and other scholars working directly on the documents cautioned that this reconstructed text would be highly unreliable for research and even questioned the ethics of using other people's work on the concordance to produce such texts. They also said they were speeding up their own research and that the official versions of these scrolls should be available by the end of the decade.

But the researchers said comparisons of some previously published sections with the same sections generated by the computer indicated the computer version is remarkably reliable.

The material the society issued Wednesday is the first of at least five volumes in "A Preliminary Edition of the Unpublished **Dead Sea Scrolls**." The volume is a transcription of historical accounts and religious laws and practices and calendars of religious holidays presumably written by monks of the small Jewish sect known as the Essenes at about the time of Jesus. The reconstructed account is in the original Hebrew.

"As far as we know, no ancient text has ever been reconstructed by computer," said Dr. Ben-Zion Wacholder, professor of Talmudic studies at Hebrew Union College and one of the researchers who compiled the reconstructed version. He was assisted by Martin Abegg Jr., a graduate student and a computer specialist.

But Dr. Frank Cross, a scholar at Harvard Divinity School who has worked with the scrolls since the 1950s, said in a telephone interview that the publication of these unauthorized versions, which he described as "pirated," would have no effect on the pace and publication schedules involving the actual scrolls.



He defended his colleagues from the frequent charges of undue secrecy and procrastination, mentioning the difficulties of working with documents that are mostly fragile fragments of parchment.

CAPTION:

PHOTO

Photo by The Associated Press

Copyright (c) 1991, Phoenix Newspapers Inc.